Verification, Validation, and Predictive Capability in Computational Engineering and Physics

William L. Oberkampf
and
Timothy G. Trucano
Sandia National Laboratories
Albuquerque, New Mexico 87185-0828

and

Charles Hirsch
Vrije Universiteit Brussel
Brussels, Belgium

Invited Presentation for:
Foundations for Verification and Validation in the 21st Century Workshop
Johns Hopkins University/Applied Physics Laboratory
Laurel, Maryland
October 22-23, 2002

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL8500

Outline of the Presentation

Engineering Sciences Center

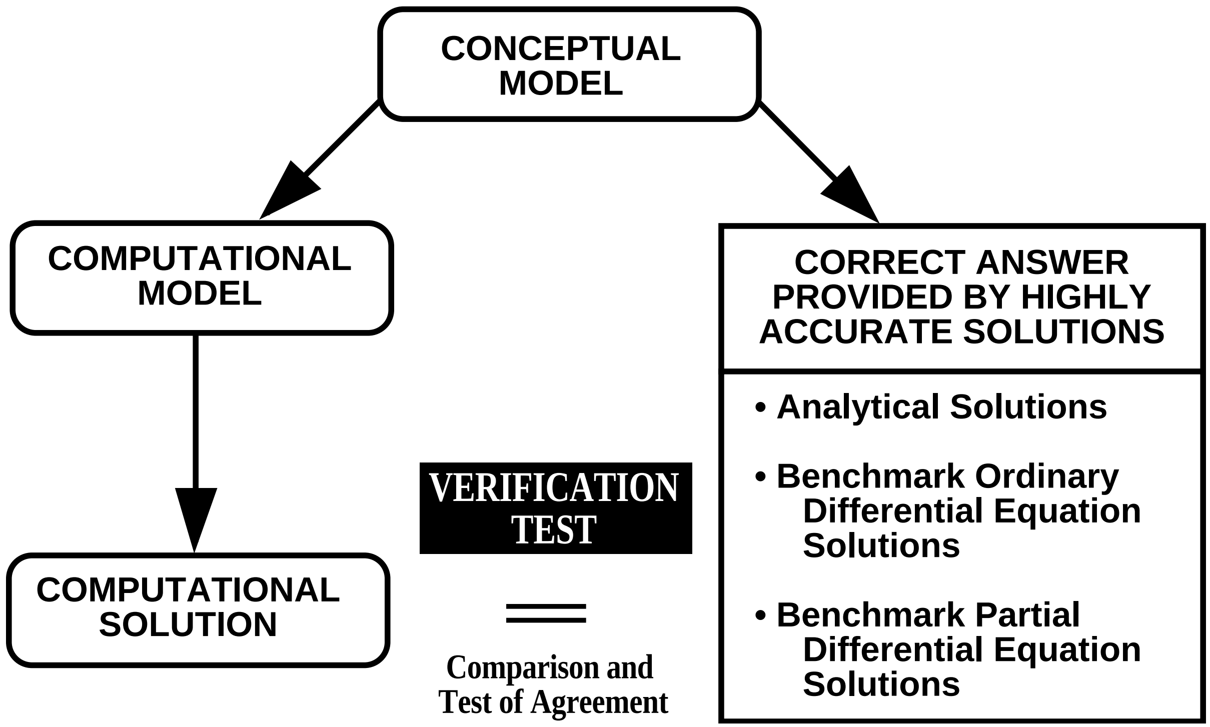
- Introduction and basic terminology
- Relationships between validation and predictive capability
- Development of verification and validation requirements
- Verification activities
- Validation activities
- Predictive capability
- Major research issues
- Major implementation issues
- Concluding remarks

Introduction

- Computational simulations have become a key contributor to:
  - Design and virtual prototyping of engineered systems
  - Supplement physical testing with virtual testing of engineered systems
  - Acquisition of new military systems
  - Certification of the performance, safety, and reliability of high-consequence systems that cannot be tested
- Why are verification and validation (V&V) important?
  - V&V are the primary means of assessing accuracy in computational simulations.
  - V&V are quantitative confidence assessment tools for computational simulations.

AIAA/DOD Definition of Verification

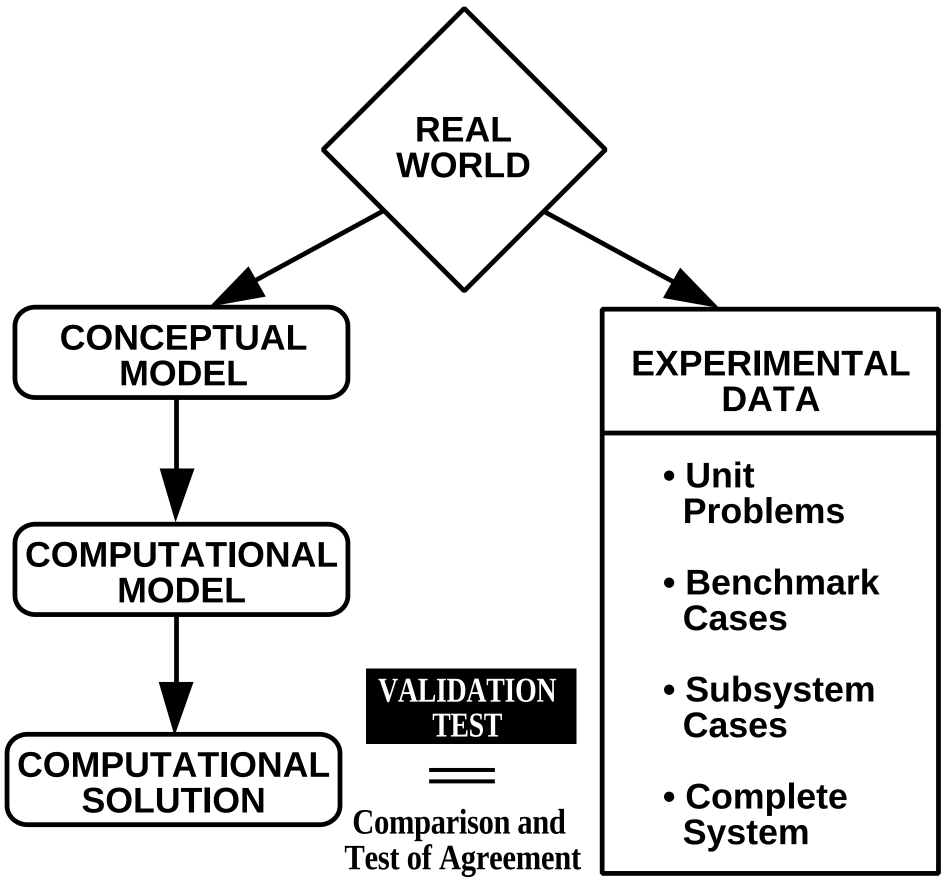
Verification: **The process of determining that a model implementation accurately represents the developer's conceptual description of the model and the solution to the model.**



DOD Definition of Validation

Engineering Sciences Center

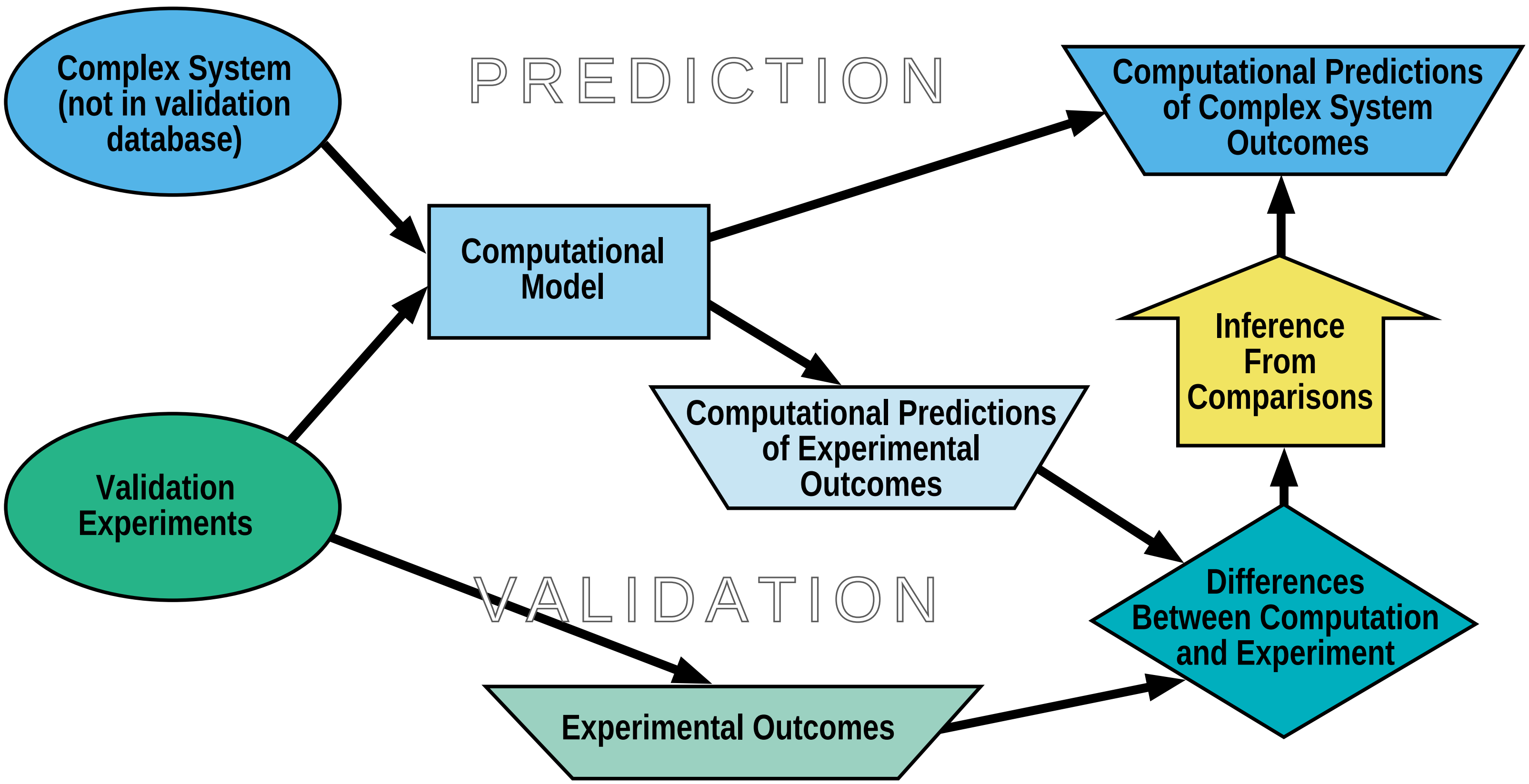
Validation: The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.



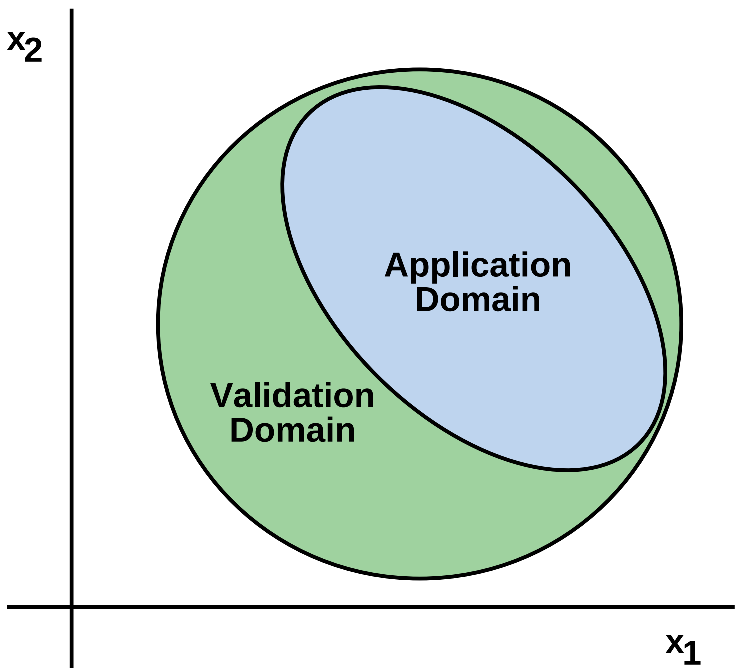
Relationships Between Validation and Predictive Capability

- **Model updating based on previous experimental data (sometimes referred to as “model validation”):**
  - **Properly referred to as parameter identification and model calibration.**
  - **Can use techniques such as Bayesian updating and Markov Chain Monte Carlo methods to determine model parameters.**
  - **Effective for engineered systems “close” to validation experiments.**
- **Alternative approach:**
  - **Appropriate for engineered systems that must operate far from the conditions under which they were validated (or calibrated).**
  - **Requires increased independence between validation and prediction.**
  - **Requires improved quantification of comparisons of computations and experiments: validation metrics**

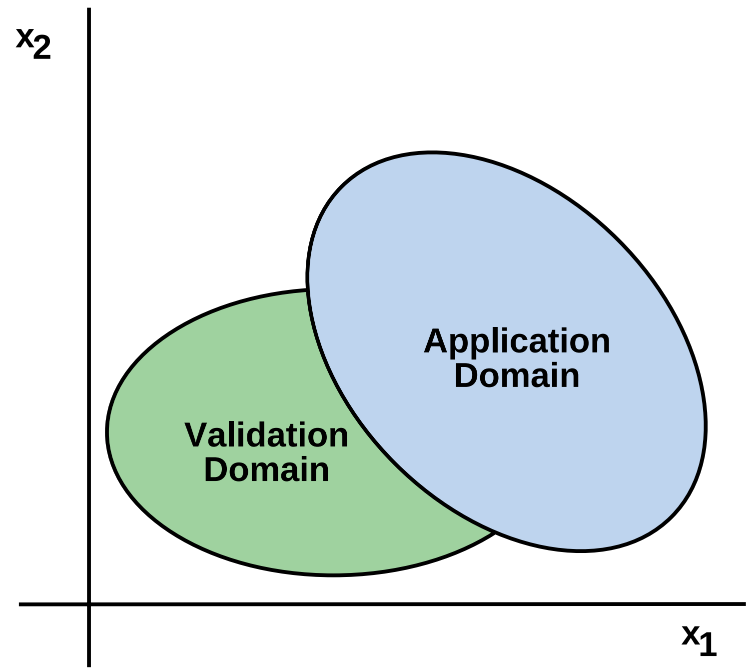
Alternative Relationship of Validation to Prediction



Common Relationships Between Validation Domain and Application Domain



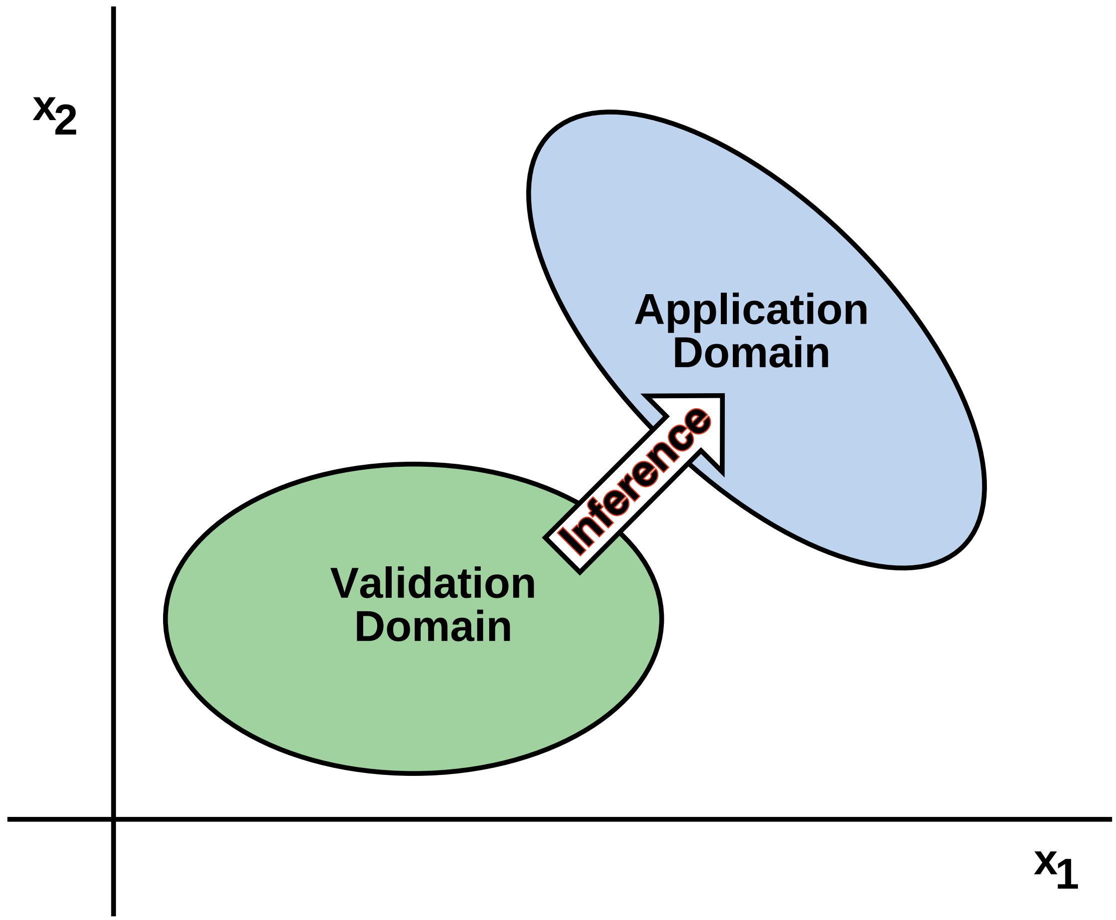
a) Complete Overlap



b) Partial Overlap

Possible Relationship Between Validation Domain and Application Domain

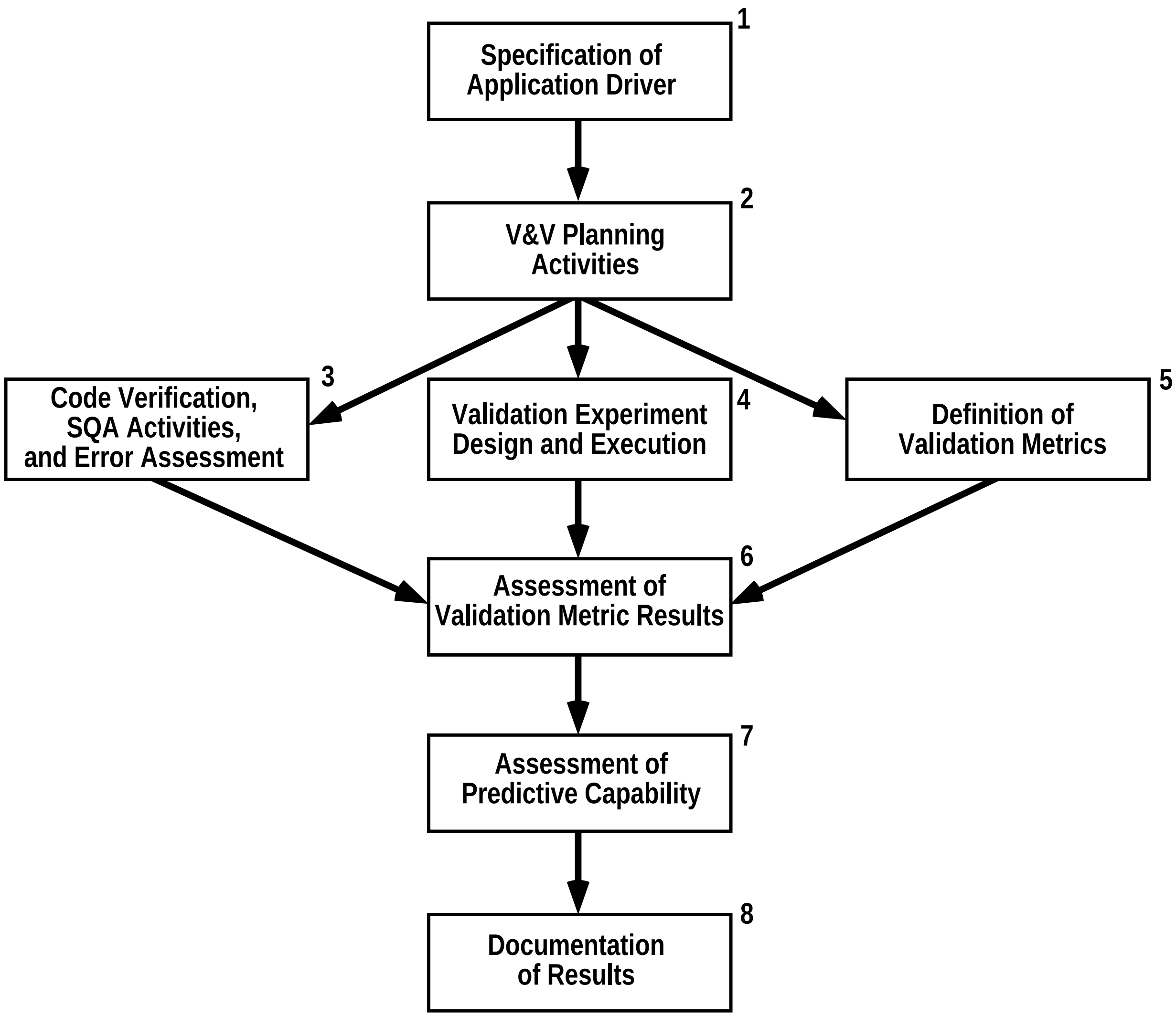
Engineering Sciences Center



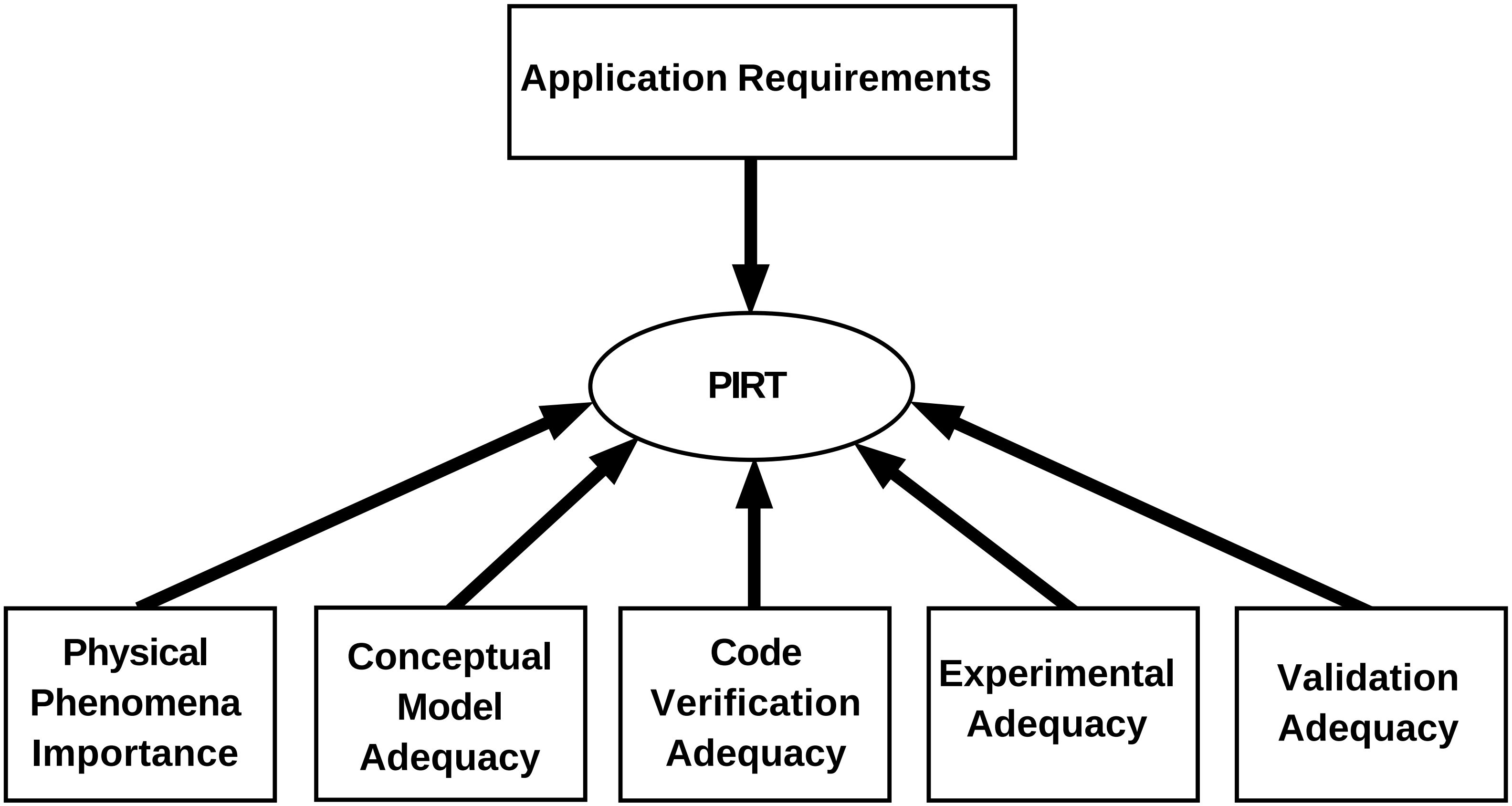
c) No Overlap

Connecting Application Requirements to V&V Requirements

Engineering Sciences Center



Using PIRT to Determine V&V Requirements and Priorities



Catagories of PIRT Information

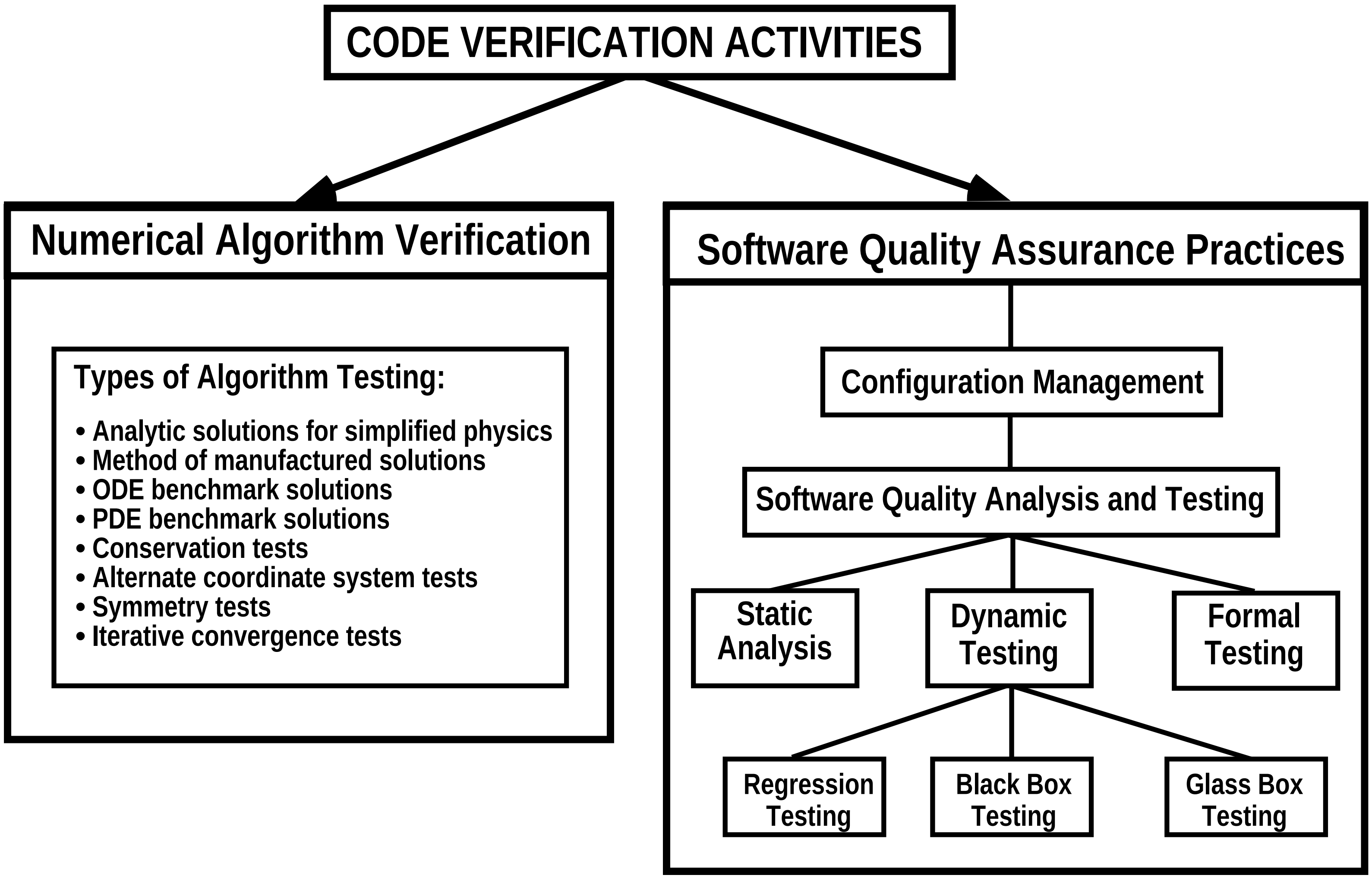
Status of Verification in Computational Physics

- **Hatton studied over 100 scientific production codes over several years. His conclusion:**

  "Scientific calculations should be treated with the same measure of disbelief researchers have for unconfirmed physical experiments."

- **We recommend that verification activities should be divided into three areas:**
  - **Numerical algorithm verification (code verification)**
  - **Software quality assurance (SQA, SQE)**
  - **Numerical error estimation (solution verification)**
- **The goals and tools of each are significantly different.**
- **However, all verification activities should deal only with the observed, i.e., *a posteriori*, performance of the code.**

Integrated View of Verification Assessment



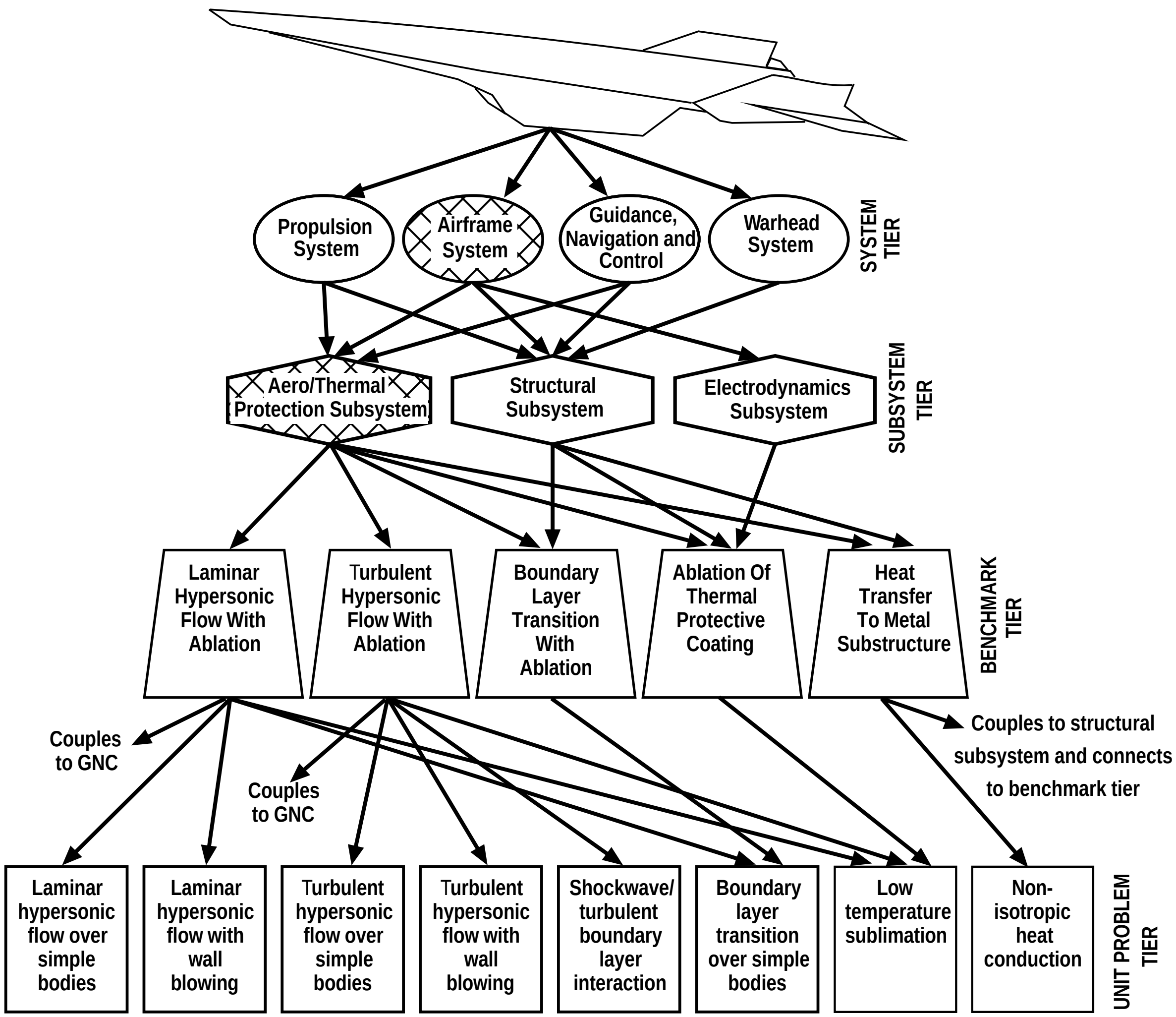
Numerical Error Estimation

- **Insufficient grid and time-step convergence is typically the largest contributor to computational error.**
- ***A posteriori*** **(vs.** ***a priori*****) methods are the only useful methods for estimating error on nonlinear partial differential equations (PDEs).**
- **Two types of grid and time-step error estimation methods:**
  - **Comparison of numerical solutions of the discretized equations on different grid sizes (related to h-adaptivity)**
  - **Comparison of numerical solutions from different discretization methods on the same grid (related to p-adaptivity)**
- **Advantages and disadvantages of each approach:**
  - **Multiple grid solutions are the most accurate and reliable**
  - **Multiple order methods require much less computational effort**

Validation Experiment Methodology: Construction of a Validation Hierarchy

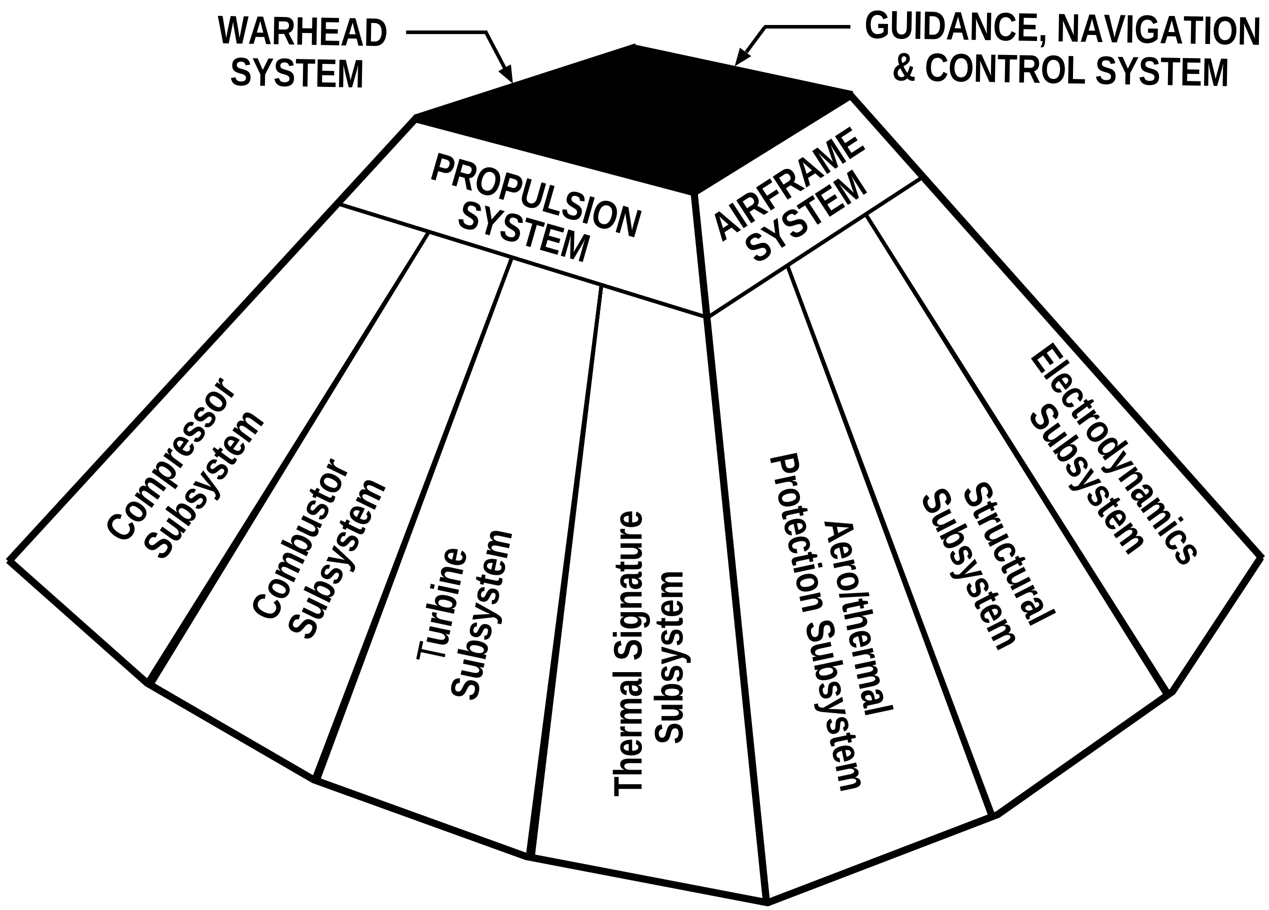
- **Validation hierarchy construction should:**
  - **Carefully disassemble the complete system**
  - **Identify experiments that are attainable and practical**
  - **Identify experiments where validation quality characterization and measurement data can be obtained**
  - **The top of the hierarchy focuses on the application of interest**
  - **The bottom of the hierarchy focuses on separate-effects physics**
- **Phenomena Identification Ranking Table (PIRT) can be used to prioritize individual validation experiments within the hierarchy.**
- **Example:**
  - **Air-launched, air breathing, hypersonic cruise missile**

Hypersonic Cruise Missile



Validation Pyramid for Hypersonic Cruise Missile

Engineering Sciences Center



Characteristics of a Validation Experiment

1. **A validation experiment should be jointly designed and executed by experimentalists and computationalists.**

2. **A validation experiment should be designed to capture the relevant physics, all initial and boundary conditions, and auxiliary data.**

3. **A validation experiment should use and develop all possible synergisms between experimental and computational approaches.**

4. **Independence between computational and experimental results should be maintained where possible.**

5. **A hierarchy of experimental measurements should be made that presents an increasing range of computational difficulty.**

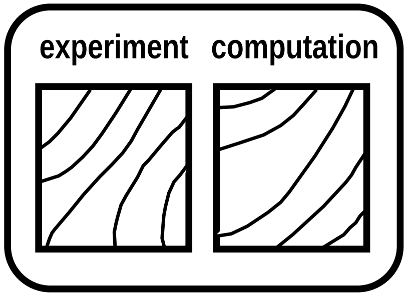
6. **Develop and employ experimental uncertainty analysis procedures to delineate and quantify random and bias errors.**

Validation Quantification

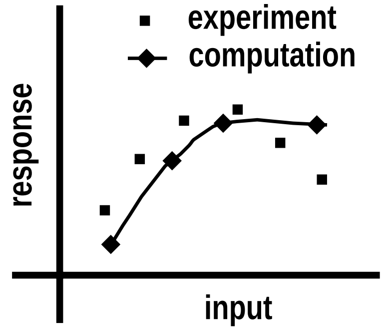
- **Approaches to validation quantification:**
  - **Model updating**
  - **Hypothesis testing**
  - **Comparison of computation and experiment**
- **Each approach relies on statistical measures because of:**
  - **Random experimental measurement error**
  - **Uncontrolled experimental parameters needed as input for computational simulations**
  - **Unmeasured experimental parameters needed as input for computational simulations**
- **We refer to the comparison of computation and experiment as a validation metric.**

Increasing Quality of Validation Metrics

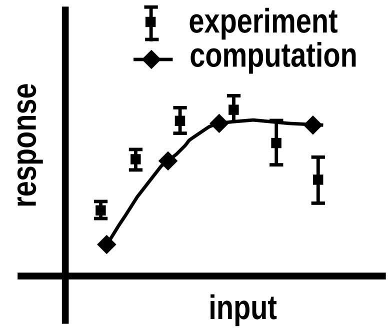
Engineering Sciences Center



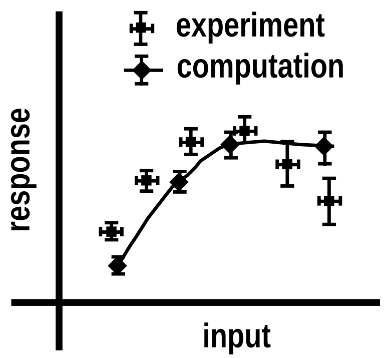
(a) Viewgraph Norm



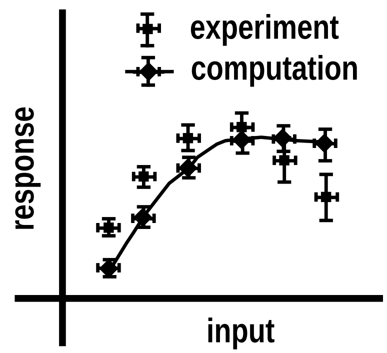
(b) Deterministic



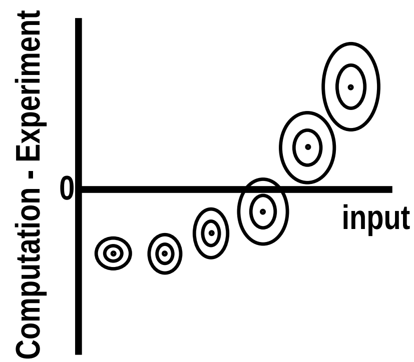
(c) Experimental Uncertainty



(d) Numerical Error



(e) Nondeterministic Computation



(f) Quantitative Comparison

Recommended Features for Validation Metrics

- **Features that should be included in a validation metric:**
  - **Should include an estimate of the numerical error.**
  - **Should include an estimate of the experimental random errors and the correlated bias errors.**
  - **Should include a test of the modeling assumptions.**
  - **Should only provide a measure of agreement between computation and experiment (not a measure of adequacy for future applications).**
  - **Should depend on the number of experimental replications of a given experimental quantity.**
  - **Should include uncertainty due to lack of experimental measurement of needed computational quantities and random uncertainty in experimental parameters.**

Predictive Capability: Uncertainty and Error

- **Uncertainty and error are commonly used interchangeably in computational physics.**
- **A number of researchers have argued that uncertainty and error should be clearly separated.**

Error: A recognizable deficiency in any phase or activity of modeling and simulation that is not due to lack of knowledge.

- **Acknowledged errors can be estimated, bounded, or ordered.**
  (discretization error, iterative error, geometry approximations)
- **Unacknowledged errors are mistakes or blunders.**
  (source code errors, compiler errors, incorrect input or output files)

Predictive Capability: Uncertainty

Engineering Sciences Center

Uncertainty: **A potential deficiency in any phase or activity of modeling process that is due to lack of knowledge.**

- **Risk assessment and information theory communities segregate uncertainty into:**
  - **Aleatory uncertainty (variability, randomness, irreducible uncertainty)**
  - **Ex: Random variation in thermodynamic properties, joint stiffness and damping due to manufacturing variability**
  - **Epistemic uncertainty (lack of knowledge uncertainty, model form uncertainty, reducible uncertainty)**
  - **Ex: poor understanding of turbulent-reacting flow and fracture dynamics, and lack of knowledge of deeply buried target characteristics**

Uncertainty and Error in Computational Predictions

- **Sources of uncertainty in computational predictions:**
  - **Validation metric will (hopefully) be a statistical measure**
  - **Conditions for the prediction will typically involve both aleatory and epistemic uncertainty**
  - **Alternate plausible models of the physical process**
- **Quantities needed for input to the predictions:**
  - **Input parameters (e.g., material properties, transport properties)**
  - **Initial conditions (e.g., initial fluid temperature, bolt preloads)**
  - **Boundary conditions (e.g., inflow conditions, forcing function)**
- **These uncertainties are normally mixed with numerical errors.**
- **Nondeterministic simulations are required to construct an ensemble of computations for a prediction.**

Three Steps for Nondeterministic Simulations

1) For the unknown or experimentally uncontrolled value construct a probability distribution.

   - Probability distributions are either determined experimentally, or assumed.

Characterizing the source of the uncertainty

2) Select input values using statistical sampling procedures.

   - Monte Carlo or Latin Hypercube procedures are typically used.

Uncertainty propagation through the computational model

3) From multiple individual computations, construct probability distributions of the required output quantities.

   - Multiple computational realizations are statistically compared with the experimentally measured quantities.

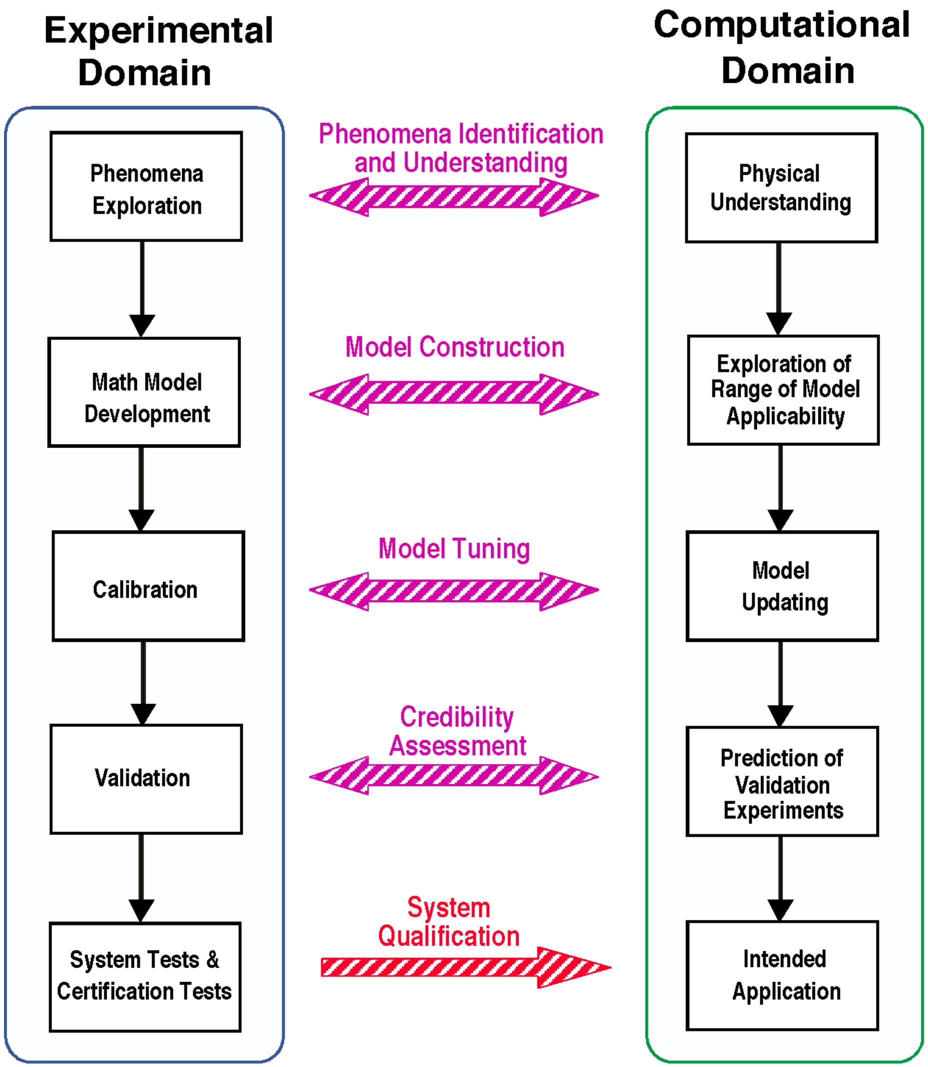
Uncertainty quantification in the computational result

Major Research Issues: Prioritization of Assessment Activities

- **The Phenomena Identification and Ranking Table (PIRT) has proven to be the most effective method for prioritization of V&V activities.**
- **PIRT can incorporate relative importance and adequacy:**
  - **Importance of multiple phenomena**
  - **Importance of multiple applications**
  - **Adequacy of alternative conceptual models**
  - **Adequacy of verification activities**
  - **Adequacy of validation simulations**
  - **Adequacy of validation experiments**
- **Methods for combining relative importance and adequacy need to be developed and evaluated.**

Major Research Issues: Prioritization of Assessment Activities

Engineering Sciences Center



Major Research Issues: Verification Activities

Engineering Sciences Center

- **The Method of Manufactured Solutions (MMS) should be developed more broadly:**
  - **Within the disciplines where it is already used, e.g., shock waves, multiphase flow, free-surface flows, and large-eddy simulation**
  - **Across more disciplines, e.g., large plastic deformation, fracture dynamics, radiation transport, and electromagnetics**
  - **Proper treatment of boundary conditions for mixed elliptic, parabolic, and hyperbolic PDEs**
- **Development of verification methods for non-unique solutions of nonlinear PDEs:**
  - **Solution bifurcation of elliptic PDEs**
  - **Chaotic solutions of hyperbolic PDEs**

Major Research Issues: Validation Activities

- **Further development and use is needed of the validation hierarchy and the system validation pyramid.**
- **Development of methods is needed to assess the adequacy, relative to the intended application, of:**
  - **Formulation of the validation metric itself**
  - **Numerical value of the metric**
- **Both tasks will be difficult because application requirements for metrics are:**
  - **Commonly not known or firm at the higher levels of the validation hierarchy**
  - **Rarely known for lower levels in the validation hierarchy**

Major Research Issues: Validation Activities (continued)

- **Additional research is needed in the formulation of validation metrics:**
  - **For steady-state problems, metrics constructed over 2D and 3D fields**
  - **For unsteady problems, formulate metrics in time**
  - **For unsteady problems with eigen frequencies, formulation of metric in the frequency domain**
  - **Methods for propagating metrics at lower levels of the validation hierarchy to higher levels of the hierarchy**
- **Bayesian updating and other calibration methods should provide additional ideas for formulation of metrics.**
- **Research is needed into determining measures of “distance” between a validation experiment and an application condition.**

Major Implementation Issues: Management Issues

- An adversarial or competitive relationship may exist between computationalists and experimentalists, either within or between organizations.
- Management must:
  - Become aware of adversarial or competitive relationships
  - Avoid any inadvertent endorsement of adversarial or competitive relationships
  - Promote synergistic relationships
- Improved methods should be found for presenting concise quantitative measures of V&V maturity.
- Verification and validation dial-meters for codes and calculations show significant promise to show relative status.

Major Implementation Issues: Management Issues (continued)

- **Management must recognize the importance of nondeterministic simulations for validation metrics and in predictive capability**
- **Systems that have heavily relied on computational simulation for safety certification requirements have fully accepted this approach:**
  - **Nuclear reactor safety**
  - **Underground storage of nuclear wastes**
- **Management must find ways to emphasize or quantify the value added by V&V activities.**
- **Factors that could be used to emphasize/quantify value added by V&V:**
  - **Professional risk to the code user of software inaccuracy/failure**
  - **Organizational risk of software inaccuracy/failure, e.g., company liability cost, environmental damage, and national security impact**

Major Implementation Issues: Industrial Setting

- **Difficulties of commercial code validation in an industrial setting:**
  - **Range of accuracy requirements between customers is very large**
  - **Split responsibility for validation between industrial users and commercial code company**
- **Industrial users must commonly deal with a complex mixture of errors and uncertainties:**
  - **Under-resolved grids**
  - **Poor grid quality**
  - **Inadequate iterative convergence**
  - **Calibration of physical model parameters**
- **Industrial users attempt to manage uncertainties and rely on the computation of incremental changes from their databases: use of deltas**

Major Implementation Issues: Commercial Software

- **Difficulties of implementing V&V procedures for commercial codes:**
  - **Code verification and SQA activities for a very wide range of computer hardware, system software, and complier software.**
  - **V&V cannot be completely tested for all possible combinations of input and output options, and internal options available in commercial software.**
  - **Bug finding, fixing, tracking, and reporting is much more difficult**
- **Documentation and availability of V&V activities of commercial software has been, in general, very poor.**
- **Electronic documentation, either in the commercial software or on the web site of the commercial software company, is recommended.**
- **User training and support of commercial software companies should be the model for industrial developed software.**

Major Implementation Issues: Development of Standards

- **Attempts should be made to standardize the meaning of validation:**
  - **Validation does not imply adequacy for applications: fidelity assessment**
  - **Validation implies adequacy for specified applications: DOD view-point**
- **Need for industry-wide standards for V&V terminology, procedures, and tools:**
  - **Professional engineering societies**
  - **Important role of European Research Community on Flow, Turbu-lence, and Combustion (ERCOFTAC) and National Agency for Finite Element Methods and Standards (NAFEMS)**
  - **These efforts should be discipline specific and should be com-posed of a very broad constituency.**

Major Implementation Issues: Development of Standards (continued)

- **Compilation, generation and documentation of highly accurate solutions for code verification should be initiated:**
  - **Sometimes referred to as strong-sense benchmarks**
  - **Role for professional engineering societies, academic institutions, nonprofit organizations, and commercial software companies.**
- **Compilation, generation and documentation of validation databases should be initiated:**
  - **Attempts have been made by the AIAA CFD Committee on Standards and National Project for Applications-oriented Research (NPARC) in CFD**
  - **Thematic Network on Quality and Trust for the Industrial Applications of CFD (QNET-CFD) is has 40 participants from several countries**
  - **QNET is funded by the European Commission**

Closing Remarks

- **Code verification on scientific and engineering software is in a dismal state, based on the comprehensive study of Hatton.**
- **Prioritization of V&V activities must be done.**
- **Validation experiments are significantly different than traditional experiments: The code is the customer**
- **Much of the existing validation experimental data will prove to be inadequate for quantitative validation**
- **New validation experiments will be costly and they will present risks to experimental facilities**
- **Validation experiments must have close cooperation between experimentalists and computationalists.**
- **The primary goal in validation is assessing the accuracy of models: not fixing or improving models.**

Closing Remarks (continued)

- **Users of the codes and users of the results of codes should require detailed documentation of V&V activities.**
- **We must find ways of convincing commercial software companies to share information on the V&V procedures they use.**
- **Industrial settings commonly require mixtures of large numerical errors and large modeling uncertainties: calibration of models must be done with extreme care.**
- **V&V activities are commonly supported as long as they:**
  - **Don't cost the code development project too much money**
  - **Don't negatively impact the code development schedule**
- **V&V activities provide short term and long term value added.**
- **We do not believe a "V&V inspection approach" will be effective.**